AI Challenge Problems
USAF-MIT AI Accelerator

Vijay Gadepally
Jeremy Kepner, Sanjeev Mohindra
vijayg@ll.mit.edu

Massachusetts Institute of Technology

Research was sponsored by the United States Air Force Research Laboratory and the United States Air Force Artificial Intelligence Accelerator and was accomplished under Cooperative Agreement Number FA8750-19-2-1000. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the United States Air Force or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

USAF-MIT AI Accelerator: Bringing world-class research to Air Force Missions

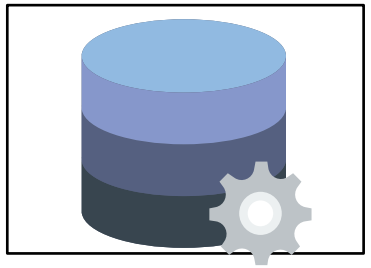
AI Researchers

MIT Campus

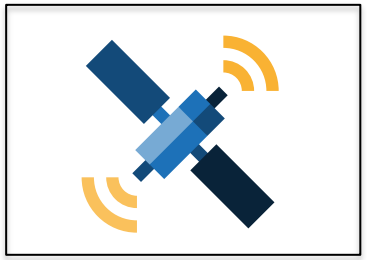
MIT Lincoln Laboratory

US Air Force

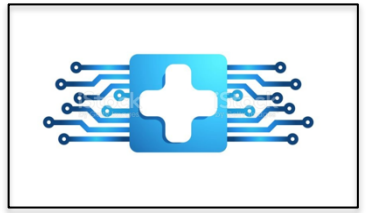
Example Air Force Missions



Computing and Cyber

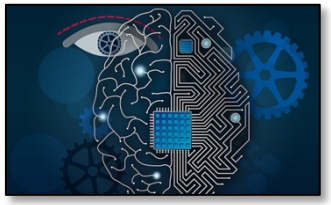


ISR and AI Assistants



Healthcare and Disaster Response

10 Mission-ready Projects



Foundational AI Research

AI Challenge Problems

AI Education and Training

USAF-MIT AI Accelerator

- Foundational
- Foundational /Applications
- Applications
- Systems

AI Accelerator (AIA) Research Projects

Foundational
Guardian Autonomy for Safe Decision Making
Foundational/Applications
Transferring Multi-Robot Learning through Virtual and Augmented Reality for Rapid Disaster Response
The Earth Intelligence
Robust Neural Differential Models for Navigation and Beyond
AI-Enhanced Spectral Awareness and Interference Rejection
Applications
Recommendations in Context over Multimedia
Multimodal Vision for Synthetic Aperture Radar (SAR)
AI-Assisted Optimization of Training Schedules
Objective Performance Prediction & Optimization Using Physiological and Cognitive Metrics
Systems
Fast AI: Datacenter and Edge Computing

150+ researchers across MIT & Lincoln working closely with dozens of AF stakeholders

AI Challenges are Critical to Breakthroughs in AI

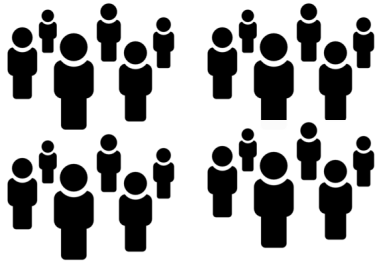
Year	Breakthroughs in AI	Datasets (First Available)	Algorithms (First Proposed)
1994	Human-level read-speech recognition	Spoken Wall Street Journal articles and other texts (1991)	Hidden Markov Model (1984)
1997	IBM Deep Blue defeated Garry Kasparov	700,000 Grandmaster chess games, aka “The Extended Book” (1991)	Negascout planning algorithm (1983)
2005	Google’s Arabic- and Chinese-to-English translation	1.8 trillion tokens from Google Web and News pages (collected in 2005)	Statistical machine translation algorithm (1988)
2011	IBM Watson became the world Jeopardy! champion	8.6 million documents from Wikipedia, Wiktionary, Wikiquote, and Project Gutenberg (updated in 2010)	Mixture-of-Experts algorithm (1991)
2014	Google’s GoogleNet object classification at near-human performance	ImageNet corpus of 1.5 million labeled images and 1,000 object categories (2010)	Convolutional neural network algorithm (1989)
2015	Google’s Deepmind achieved human parity in playing 29 Atari games by learning general control from video	Arcade Learning Environment dataset of over 50 Atari games (2013)	Q-learning algorithm (1992)
Average No. of Years to Breakthrough:		**3 years**	**18 years**

Source: Train AI 2017, https://www.crowdflower.com/train-ai/

Winning AI

1 Team	1 Problem	Win Probability
		~20%

10 Teams	10 Problems	Win Probability
		~200% (~2 solutions to 10 problems)

10 Teams	10 Problems	10 Challenges	100s of Teams	Win Probability
				>>1000% (many solutions to each problem)

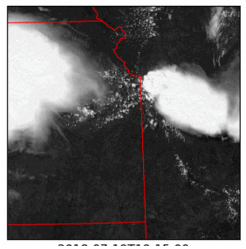
AIA Challenge Problems

- **Goal of AIA challenge problems is to drive innovation in the wider AI ecosystem**
  - **Other academic institutions**
  - **Government Laboratories**
  - **Small and medium-sized businesses**

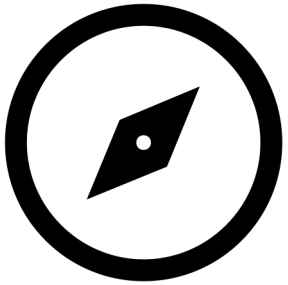
- **Each AIA team is putting together a series of challenge problems that they will make available to the wider community**
  - **Developed in collaboration with Air Force stakeholders, research collaborators**

- **Challenge problems consist of:**
  - **Relevant datasets**
  - **Mathematical specifications**
  - **Baseline implementations**
  - **Evaluation criteria**

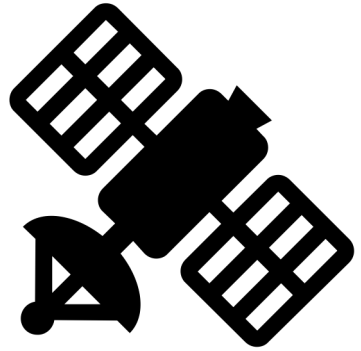
AIA Challenge Problem Examples



SEVIR Nowcasting Challenge

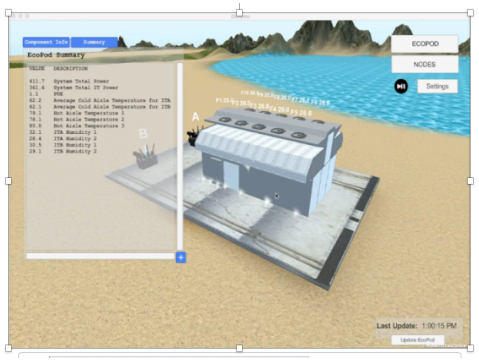


Signal Enhancement for Magnetic Navigation

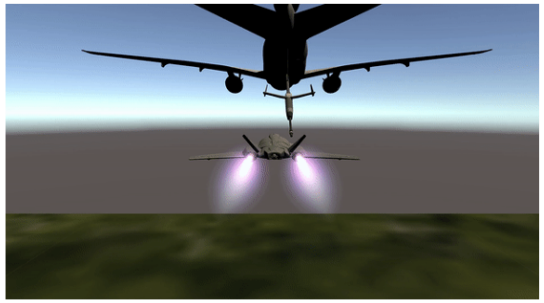


RFChallenge

Pilot Performance Assessment Challenge



Datacenter Challenge

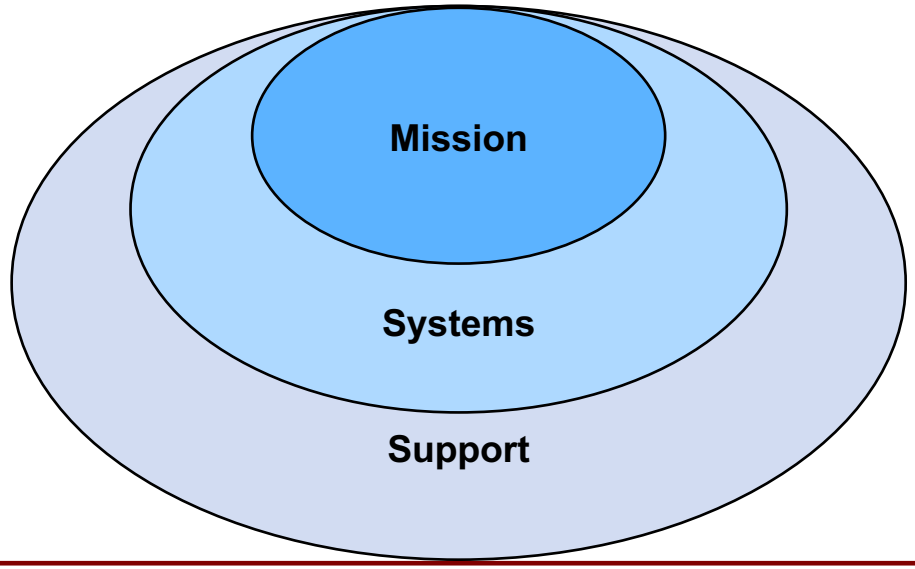


Air Force Arcade

Example AIA Challenge Problem
- Fast AI Datacenter Challenge -

Air Force Problem

- Various systems throughout the Air Force that support multiple mission objectives
  - E.g., datacenters, airframes, radars, …
- Integrating disparate data sources to look for outliers or anomalies is a major challenge
- Desired end state of proactive system maintenance through behavioral understanding to improve security, fault predictability, usage efficiency.



Challenge Name: Datacenter Challenge
Project: Fast AI

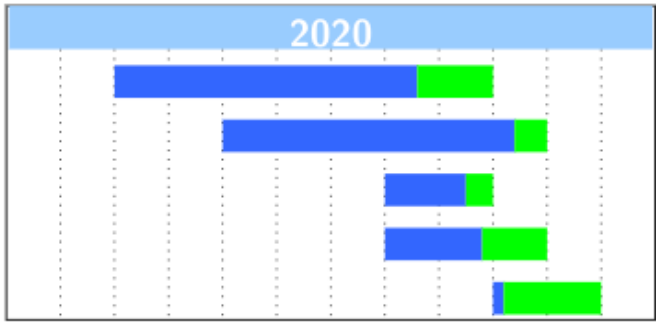
- AI challenge that enables datacenters to:
  - Predict and identify system failures from multi-modal data
  - Leverage AI to improve datacenter operations
  - Identify and stop policy violations

- Challenge Website: datacenterchallenge.mit.edu

Air Force Stakeholders

412th Test Wing

Kessel Run

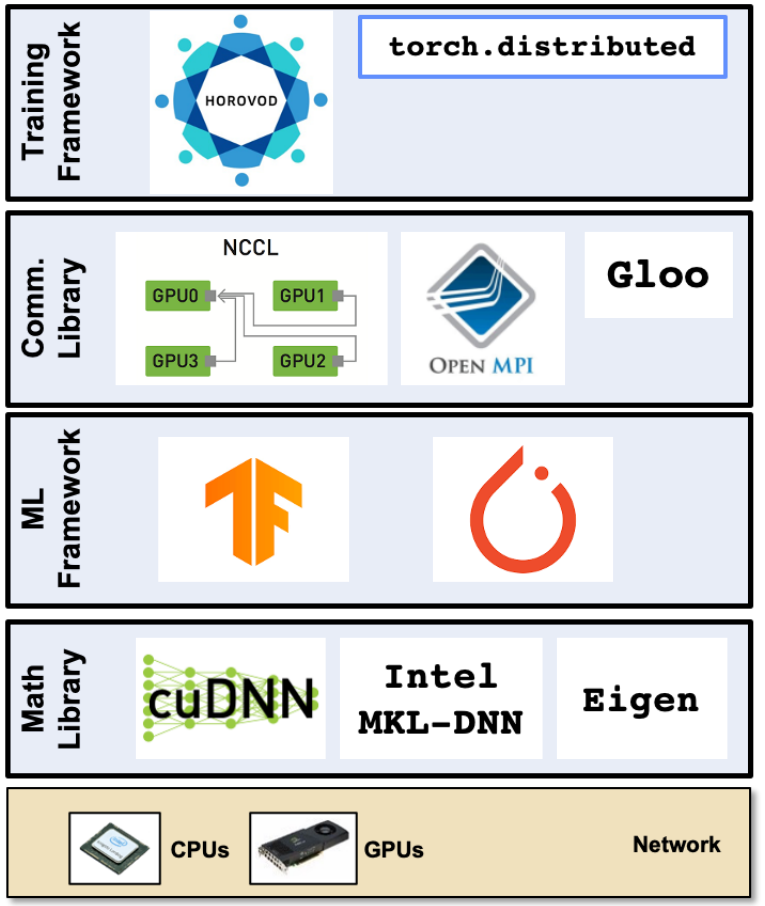


Fast AI Challenge Problem Team: Siddharth Samsi, Daniel Edelman, David Bestor, Vijay Gadepally, Benjamin Price

Datacenter Challenge: Challenge Problem to Address Air Force Needs

- AI workloads are complex and rely on a deep hardware/software stack.
- High performance computing centers increasingly support AI/ML as well as traditional compute workloads.

- Goals:
  - Optimize system scheduling for improved resource consumption
  - Suggest optimization pathways for users
  - Predict and identify system failures
  - Identify and stop policy violations



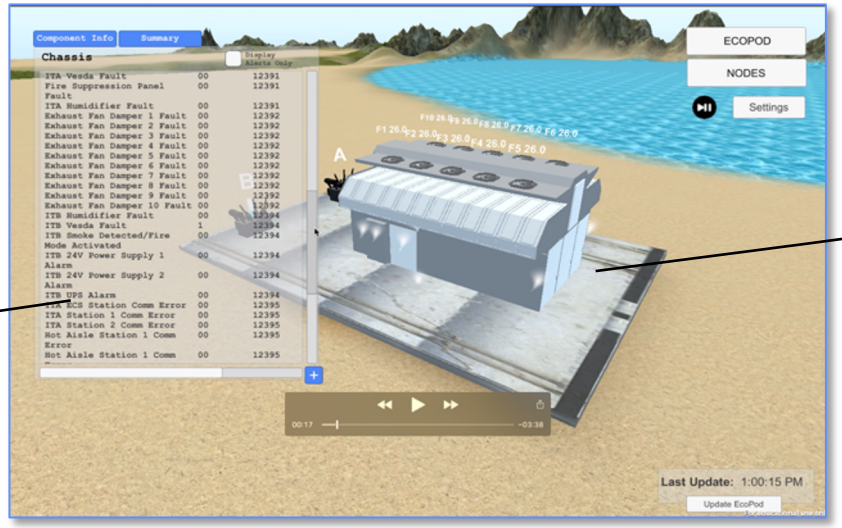
The Fast AI Datacenter Challenge aims to foster innovation in AI approaches to the analysis of large scale datacenter monitoring logs

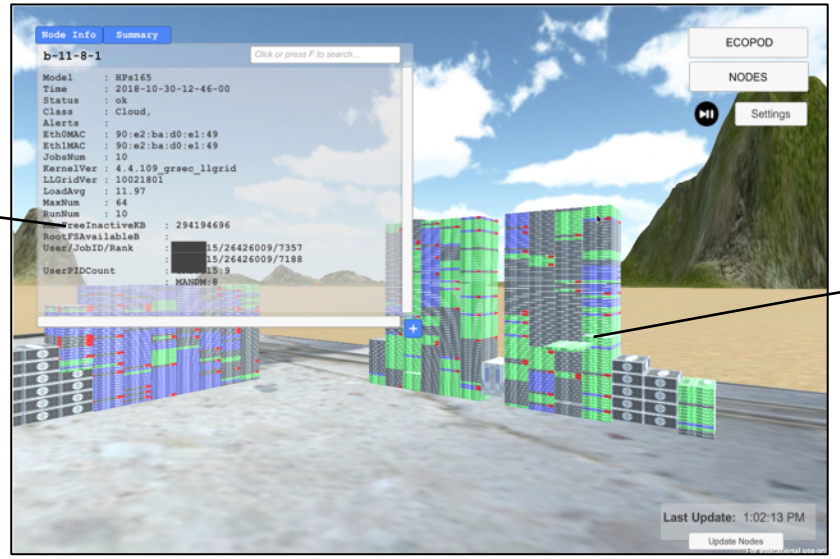
Data Collected
-From Lincoln Laboratory Supercomputing Center-

Building management:
- Electrical data
- Water detection
- Power supply alarms
- Smoke/Fire alarms
- Exhaust fan alarms

Environmental data:
- Humidity
- Temperature
- Air flows

Node level data:
- Username
- Job name
- Job status
- Job start time
- Nodes assigned to the job



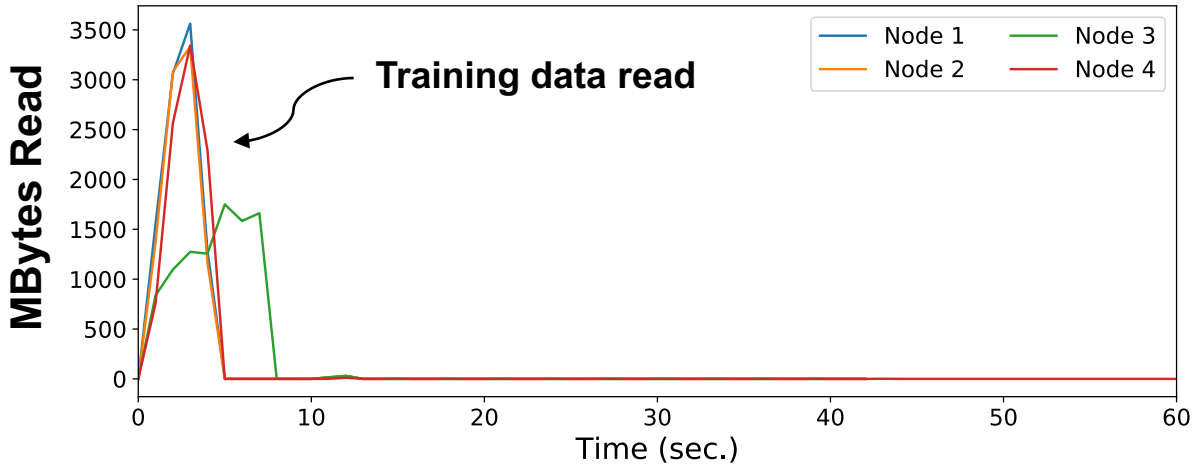


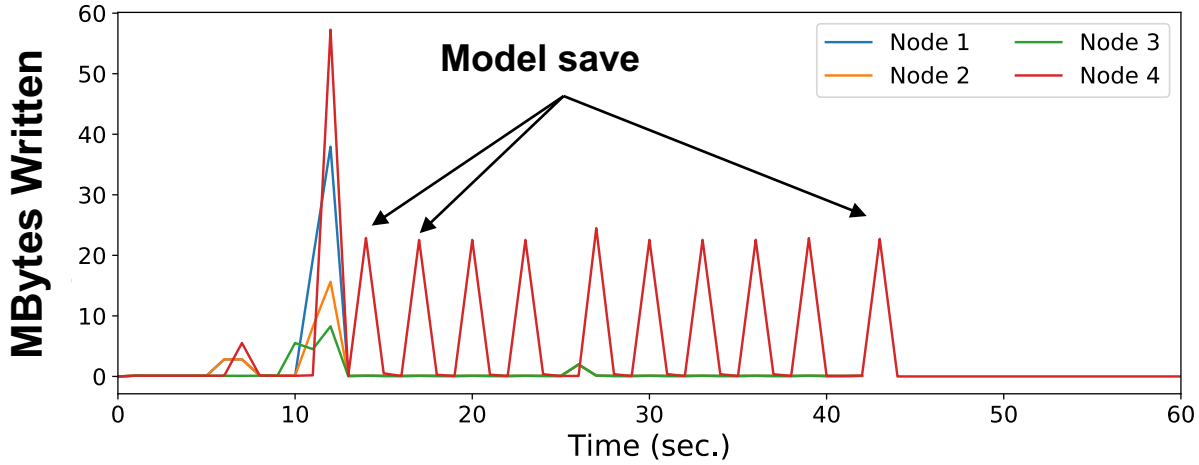
System level data:
- Number of jobs running vs queued
- Job breakdown by user
- Job breakdown by resource (GPU vs CPU)

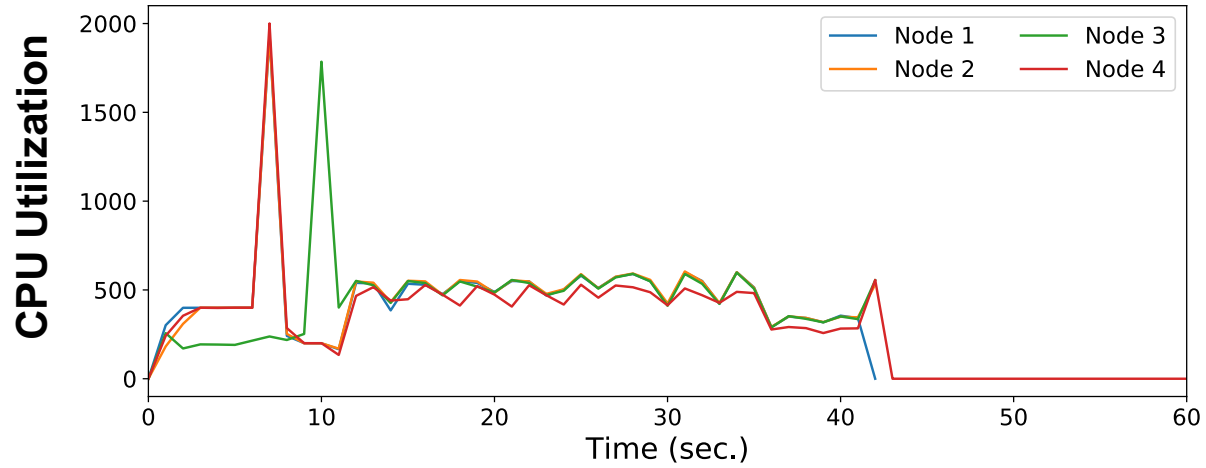
Network traffic data: *(tentative)*
- Host level `tcpdump`
- Network Flows
- Interface counters/packet corruption

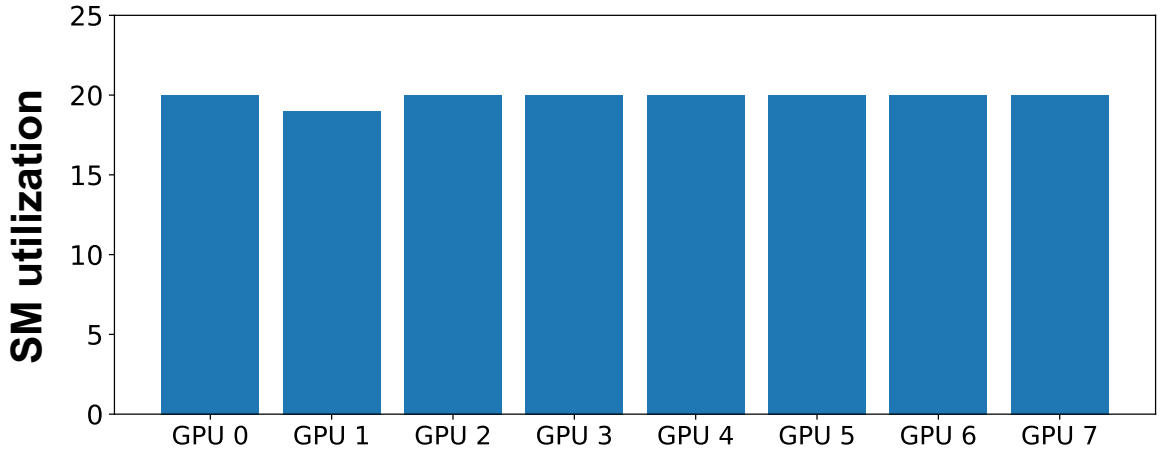
Job level data:
- Node name
- Number of processes
- MAC address of node
- Average load on the system
- Memory usage on node
- Number of Lustre calls

Data Example: Time Series Data and GPU Statistics







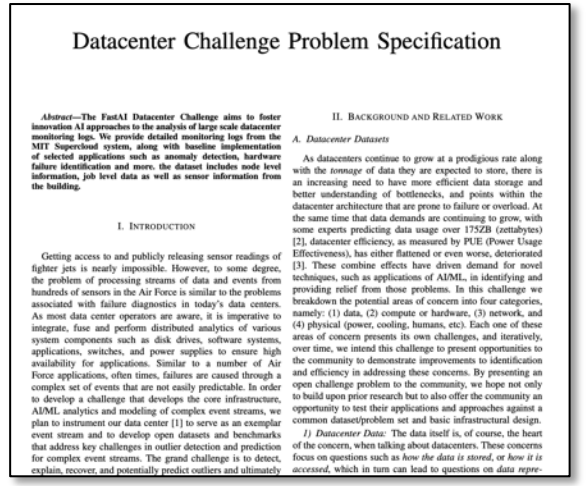


Statistics from data distributed training on 4 nodes and 8 GPUs

SM : Streaming multiprocessors

Datacenter Challenge: Dissemination

- **Challenge specification and links to dataset will be available on https://datacenterchallenge.mit.edu**



Datacenter Challenge Problem Specification

Abstract—**The FastAI Datacenter Challenge aims to foster innovation AI approaches to the analysis of large scale datacenter monitoring logs. We provide detailed monitoring logs from the MIT Supercloud system, along with baseline implementation of selected applications such as anomaly detection, hardware failure identification and more. the dataset includes node level information, job level data as well as sensor information from the building.**

I. Introduction

Getting access to and publicly releasing sensor readings of fighter jets is nearly impossible. However, to some degree, the problem of processing streams of data and events from hundreds of sensors in the Air Force is similar to the problems associated with failure diagnostics in today's data centers. As most data center operators are aware, it is imperative to integrate, fuse and perform distributed analytics of various system components such as disk drives, software systems, applications, switches, and power supplies to ensure high availability for applications. Similar to a number of Air Force applications, often times, failures are caused through a complex set of events that are not easily predictable. In order to develop a challenge that develops the core infrastructure, AI/ML analytics and modeling of complex event streams, we plan to instrument our data center [1] to serve as an exemplar event stream and to develop open datasets and benchmarks that address key challenges in outlier detection and prediction for complex event streams. The grand challenge is to detect, explain, recover, and potentially predict outliers and ultimately

II. Background and Related Work

A. Datacenter Datasets

As datacenters continue to grow at a prodigious rate along with the *tonnage* of data they are expected to store, there is an increasing need to have more efficient data storage and better understanding of bottlenecks, and points within the datacenter architecture that are prone to failure or overload. At the same time that data demands are continuing to grow, with some experts predicting data usage over 175ZB (zettabytes) [2], datacenter efficiency, as measured by PUE (Power Usage Effectiveness), has either flattened or even worse, deteriorated [3]. These combine effects have driven demand for novel techniques, such as applications of AI/ML, in identifying and providing relief from those problems. In this challenge we breakdown the potential areas of concern into four categories, namely: (1) data, (2) compute or hardware, (3) network, and (4) physical (power, cooling, humans, etc). Each one of these areas of concern presents its own challenges, and iteratively, over time, we intend this challenge to present opportunities to the community to demonstrate improvements to identification and efficiency in addressing these concerns. By presenting an open challenge problem to the community, we hope not only to build upon prior research but to also offer the community an opportunity to test their applications and approaches against a common dataset/problem set and basic infrastructural design.

1) Datacenter Data: The data itself is, of course, the heart of the concern, when talking about datacenters. These concerns focus on questions such as *how the data is stored*, or *how it is accessed*, which in turn can lead to questions on *data repre-*

- **Initial announcement at IEEE HPEC 2020**
  - **Formal challenge specification underway**

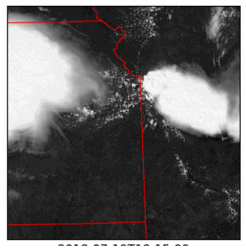
- **Other dissemination venues being considered: Workshops/BoFs at conferences such as IEEE/ACM Supercomputing, IPDPS and others**

SC20 Atlanta, GA | more than hpc.

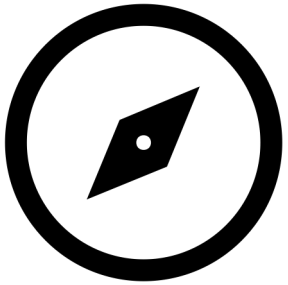
IPDPS 2021 Portland

IEEE HPEC High Performance Extreme Computing

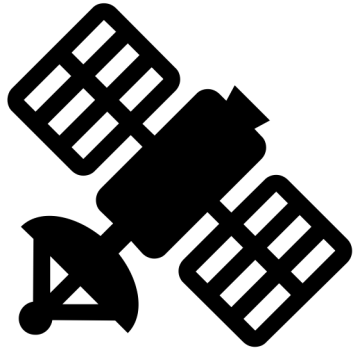
AIA Challenge Problem Examples



SEVIR Nowcasting Challenge

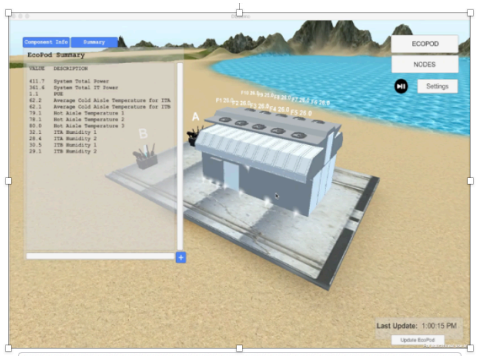


Signal Enhancement for Magnetic Navigation

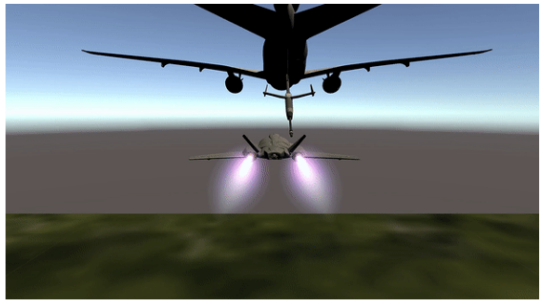


RFChallenge

Pilot Performance Assessment Challenge



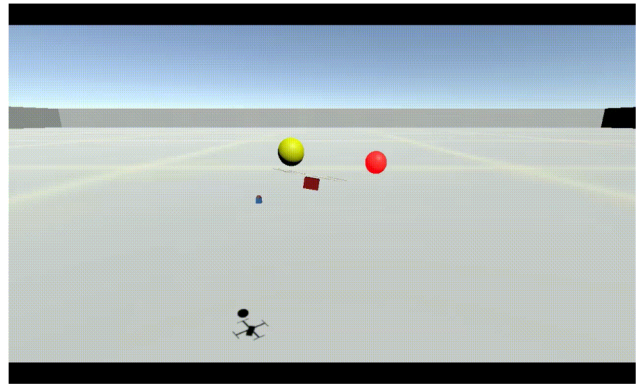
Datacenter Challenge



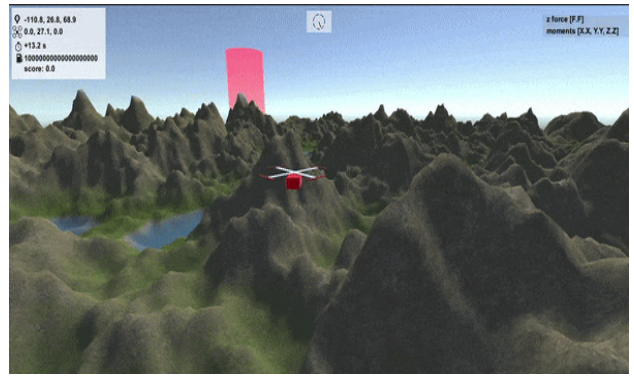
Air Force Arcade

Challenge Name: Air Force Arcade
Project: Air Guardian

- Overview of Challenge Problem
  - Use reinforcement learning and other techniques to solve Air Force-related game scenarios in continuous control space.
  - Provide RL environments to the research community that cover underrepresented control and observation schemes.



- Target challenge audience
  - Reinforcement learning and autonomous flight community



- Challenge release target: December 2020

Air Guardian Challenge Problem Team: Hosea Siu, Catherine Johnson, Kyle Palko, Ross Allen, Paul Tylkin

Summary

- US Air Force – MIT AI Accelerator project teams are developing a series of AI challenges
- Each challenge is designed to drive innovation in applying AI to different missions and problems
- Keep an eye on https://aia.mit.edu for more updates!
- Email: vijayg@ll.mit.edu